

Strategy:

1. First create varied keypoint definitions (Biological, Visible, ..)
2. Then label AP10k antelope images (around 200) with all definitions
3. Train RTMPose model on these separate datasets and compare accuracy across definitions (including the original AP10k)

Alternate strategy (to avoid labeling so much data):

1. First create varied keypoint definitions (Biological, Visible, ..)
2. Then label few AP10k antelope images with all definitions
3. Compare labeler consistency for all definitions using variance
4. Then label AP10k antelope images (around 200) with the definition that was most consistent
5. Train RTMPose model on this dataset and compare accuracy to original AP10k